Markov Random Fields (MRF)

Julien Burkhard

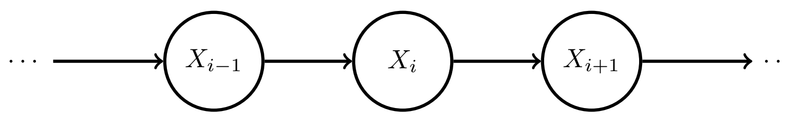
March 14, 2024

1 Markov Chains

In a first order Markov chain (MC):

$$\Pr(X_i \mid X_{i-1}, X_{i-2}, \ldots, X_1) = \Pr(X_i \mid X_{i-1})$$

For example with:



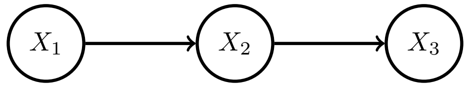
For the sake of argument, let's consider $X_i \in \mathcal{F}_i$ as the event at time i from the possible events $\mathcal{F}_i$. In all generality, we would expect for a causal process to simplify to:

$$\Pr(X_i \mid \ldots, X_{i+1}, X_{i-1}, \ldots, X_1) = \Pr(X_i \mid X_{i-1}, X_{i-2}, \ldots, X_1)$$

By Markov's property, it adds an additional constraint and says that:

$$\Pr(X_i \mid X_{i-1}, X_{i-2}, \ldots, X_1) = \Pr(X_i \mid X_{i-1})$$

In other words, it is the **conditional independence** of X_i w.r.t X_j where $i \neq j$ but not the independence !



As we can see for X_1, X_2, X_3.

$$\begin{aligned}\Pr(X_3, X_1) &= \Pr(X_3 \mid X_1)\Pr(X_1) \\ &= \Pr(X_3 \mid X_1)\Pr(X_1) \\ &= \Pr(X_1) \sum_{X_2 \in \mathcal{F}_2} \Pr(X_3 \mid X_2)\Pr(X_2 \mid X_1)\end{aligned}$$

Then independence is satisfied only if:

$$\Pr(X_3) = \sum_{X_2 \in \mathcal{F}_2} \Pr(X_3 \mid X_2)\Pr(X_2 \mid X_1)$$

Which is not necessarily the case. The following relation is used:

$$\Pr(X_3 \mid X_1) = \sum_{X_2 \in \mathcal{F}_2} \Pr(X_3 \mid X_2)\Pr(X_2 \mid X_1)$$

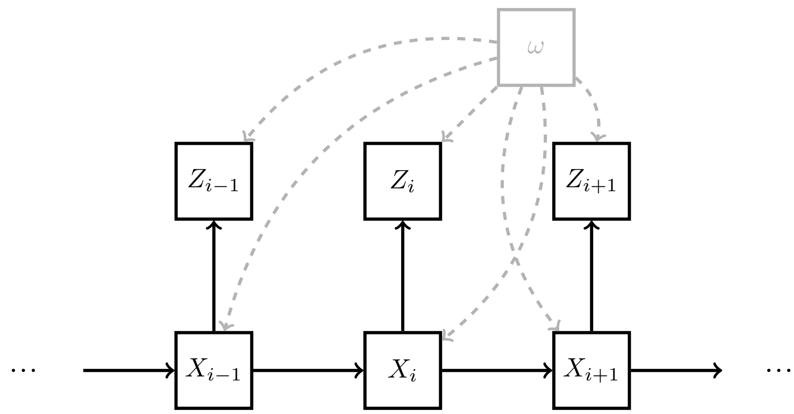
Which is useful to express the dependence events to anterior events. More generally the event at state i is dependent on the event $i - N$ with:

$$\Pr(X_i \mid X_{i-N}) = \sum_{X_{i-2} \in \mathcal{F}_{i-2}, \ldots, X_{i-N+1} \in \mathcal{F}_{i-N+1}} \Pr(X_{i-2} \mid X_{i-2}) \ldots \Pr(X_{i-N+1} \mid X_{i-N})$$

So relating to algorithms and data structures, the Markov chain formulation may be a instance of memoization to solve a Markov process joint probability !

2 Hidden Markov Model (HMM)

The problem is as follows. We are given a series of observations Z_i where there are hidden events X_i modeled by a Markov chain. We may also have parameters for the model ω. All of these are expressed through a probabilistic graph model such as this one:



Where we can express the likelihood of $\boldsymbol{Z}_N = \{Z_1, \dots, Z_N\}$ w.r.t $\boldsymbol{X}_N = \{Z_1, \dots, Z_N\}$ as:

$$\Pr(\boldsymbol{Z}_N \mid \boldsymbol{X}_N) = \prod_i^N \Pr(Z_i \mid X_i)$$

The continous version of HMM is actually a Kalman filter !

3 Markov Random Fields (MRFs)

MRFs like MCs also define a conditional independence property:

$$\Pr(X_i \mid \dots, X_{i+1}, X_{i-1}, \dots, X_1) = \Pr(X_i \mid X_{j \in \mathcal{N}(i)})$$

where $\mathcal{N}(i)$ is the set of events directly adjancent to i.

The joint probability which in MCs is computed on the last state is now defined as:

$$\Pr(\boldsymbol{X}) \propto \prod_{\boldsymbol{c} \in \mathcal{C}} \Pr(\boldsymbol{c})$$

Where $\mathcal{C}$ is the set of maximum cliques. This is the result of Hammersley-Clifford theorem.

Factor graphs makes which cliques are choosed explicit in the graph.